Customer Joining-balking Strategies in an Observable Queue with Partial Service Time Information

Wei Sun, Yan Wang, Kaibo Yuan, and Shiyong Li, *Member, IAENG*

***Abstract*—This paper concerns customers' decision process of joining or balking at their arrival instants in an observable queue. Given partial information that different number of moments of the service time is available to the arriving customers, they need to estimate its distribution by using only the information available and no other information based on the maximum entropy principle to decide whether join or balk. Respectively for the systems with risk-neutral and risk-averse customers, if the actual service time distribution is exponential, we numerically compare the customers' equilibrium threshold strategies under several types of partial information and socially optimal as well as profit-maximizing threshold strategies. We find that in general the equilibrium thresholds are no less than the socially optimal one which is also no less than the profit-maximizing one, and verify that more risk-averse customers hold lower threshold.**

***Index Terms*—Observable queue, Partial information, Moments, Equilibrium, Social optimization, Profit maximization.**

I. Introduction

So far, much excellent literature on the relationship between the delay information and customers' strategic behavior came forth. Hassin and Haviv [1] gave an excellent survey on this topic. Then Guo and Zipkin [2] considered three levels of delay information: no information, queue length, and exact waiting time. They focused on ways to compute the performance measures in the three systems and proved that social welfare need not increase when more accurate delay information is available. This work is generalized into a system with phase-type service times by Guo and Zipkin [3]. Then they [4] considered two types of vague information about delays, besides the basic three ones in Guo and Zipkin [2]. In the first one with partition information, customers learn a rough range of the current queue length, whereas in the second one with phase information, customers learn the total number of phases remaining in the system. Ibrahim and Whitt [5] explored the performance of different real-time delay estimators based on recent delay experiences of customers, and Ibrahim and Whitt [6] proposed alternative, effective ways to implement better delay estimation in overloaded, many-server queues with customer abandonment. Armony et al. [7] investigated the performance impact of making delay announcements to arriving customers in a many-server queue setting with customer abandonment. Furthermore, Jouini et al. [8] formulated a multi-class call center model with priorities and impatient customers who are announced the delay information upon their arrival. Then Chen et al.[9] also analyzed a call center with partial closing rules, feedback and impatient customers.

Recent years, Sun et al. [10][11] considered three types of setup/closedown policies and derived customers' equilibrium threshold strategies in observable queues and equilibrium mixed strategies in unobservable queues, respectively. Subsequently, Guo and Hassin [12][13] studied homogeneous and heterogeneous customers' equilibrium and socially optimal behavior in observable and unobservable queues with N-policy, respectively, then Economou et al. [14] derived the customers' optimal balking strategies in single-server queues with general service and vacation times. As for vacation queues, Sun and Li [15] compared equilibrium mixed strategies of risk-neutral and risk-averse customers with different decision criteria in some unobservable queues with multiple vacations. With respect to different levels of system information, Sun and Li [16], Sun et al. [17] and Wang et al. [18] focused on customers' equilibrium/socially optimal balking behavior in Markovian queues with working vacations. Most recently, Panda et al. [19] considered customers' equilibrium and optimal balking behavior in a single-server Markovian queue with multiple vacations and geometric abandonments. Then Sun et al. [20] focused on the same problem in unobservable queues with double adaptive working vacations.

All the above works study information on queues, such as queue length or waiting time, or server state. However, in this paper we study the information on service times. Considering very often the server doesn't convey the full information about the service time he has acquired to customers deliberately or because of objective factors, the customers needs to estimate the probability distribution of service time by using only the information available and no other information to make decision of joining or balking at their arrival instants. There exist few works studying this type of information in a queue to analyze the customers' behavior except for Guo et al. [21]. They studied customer equilibrium as well as socially optimal strategies to join an unobservable queue with only partial information on the service time distribution, such as moments and the range. However, this paper discusses customers' joining-balking behavior in an observable queue with partial information that

Manuscript received October 22, 2017; revised July 13, 2018. This work was supported in part by the National Natural Science Foundation of China (Nos. 71671159 and 71301139), the Humanity and Social Science Foundation of Ministry of Education of China (No. 16YJC630106), the China Postdoctoral Science Foundation (No. 2018T110205), the Natural Science Foundation of Hebei Province (Nos. G2016203236, G2018203302 and G2016203220), the project Funded by Hebei Education Department (Nos. BJ2017029 and BJ2016063), and the project Funded by Hebei Talents Program (No. A2017002108).

Wei Sun, Yan Wang, Kaibo Yuan and Shiyong Li are with School of Economics and Management, Yanshan University, Qinhuangdao 066004, China. Shiyong Li is the corresponding author. e-mail: shiyongli@ysu.edu.cn.

different number of raw moments of the service time.

For the estimation method, we assume the customers all conform with the maximum entropy principle. To acquire more, the beginner can consult the literature given by Kapur and Kesavan [22] where an excellent and detailed introduction about the maximum entropy principle is presented. The objective of the maximum entropy principle is to choose, under given information, the maximum entropy distribution in all possible compatible distributions. Based on the maximum entropy principle and by using Lagrange undetermined multiplicator method, we may get the maximum entropy distribution consistent with the given information. Therefore, the maximum entropy distribution of the service time gives customers the most unbiased and most objective distribution consistent with the partial information.

We first derive the equilibrium thresholds under the partial information that different number of service time moments for both the systems with risk-neutral and risk-averse customers. Especially given that the actual service time distribution is exponential, we derive the socially optimal and profit-maximizing thresholds for two risk-averse cases and then numerically compare all kinds of thresholds, and find that it is kept in general for the risk-neutral and risk-averse customers that the equilibrium threshold(s) is (are) no less than the socially optimal one(s) which is also no less than the profit-maximizing one(s). Then we summarize that the equilibrium threshold under the partial information of mean and variance is greater than that under the partial information of mean, and verify that the more risk-averse customers are, the lower threshold they hold.

In Section 2, we first consider the customers' equilibrium threshold strategies under the partial information of different number of service time moments. Given the actual service time distribution is exponential, we get the corresponding equilibrium thresholds and socially optimal one of the customers in Section 3, and then derive the profit-maximizing thresholds of the server for two risk-averse cases in Section 4. Meanwhile, we numerically compare the equilibrium and socially optimal as well as profit-maximizing thresholds for both the risk-neutral and risk-averse cases. Finally, we briefly conclude the paper and put forward some future work.

II. EQUILIBRIUM

First assume service reward R and unit waiting cost c if any customer joins, then a Poisson arrival process with parameter λ to a single server system and a tagged customer can observe the queue length $n-1$ at his arrival instant, whereas he is not told about the exact distribution function F_S of the service time S except some partial information—some first moments of the service time. Denote F as the distribution function of the customer's sojourn time $T = S_1^- + S_2 + \cdots + S_n$, where S_1^- is the residual service time of the customer who is just receiving service when the tagged customer arrives. Specially given exponentially service time, $\{S_1^-, S_k, 2 \le k \le n\}$ are all i.i.d. random variables which have the same distribution with S, that is, F is the n–fold convolution of the service time distribution function F_S.

A. Risk-Neutral Customers

First consider a system with risk-neutral customers, in which any customer's residual utility of joining U_{join}, maybe positive or negative, can be expressed as

$$\begin{aligned} U_{join} &= \int_0^\infty (R - ct)dF(t) \\ &= R - cE[T] \\ &= R - c(n-1)\overline{s}_1 - c\frac{\overline{s}_2}{2\overline{s}_1}, \end{aligned} \tag{1}$$

where $\overline{s}_1 = E[S]$ and $\overline{s}_2 = E[S^2]$. As for $E[T]$, because $E[S_1^-] = \overline{s}_2/(2\overline{s}_1)$, we have $E[T] = (n-1)\overline{s}_1 + \overline{s}_2/(2\overline{s}_1)$. Obviously, we should assume that the customer's residual utility of balking $U_{balk} = 0$. If $U_{join} > U_{balk} = 0$, then the arriving customer will join the queue. Otherwise, the arriving customer will balk.

If arriving customers only know the mean $\overline{s}_1$ of the service time, then the following lemma shows the form of the maximum entropy distribution based on the partial information (see Kapur [22]).

Lemma 1 If the service time is positive with its mean $\overline{s}_1$, then the maximum entropy distribution of the service time is exponential with parameter $1/\overline{s}_1$.

Theorem 1 If risk-neutral customers are informed of the mean $\overline{s}_1$ of the service time in an observable queue, then the equilibrium threshold

$$n_{eM} = \left\lfloor \frac{R}{c\overline{s}_1} \right\rfloor.$$

If $0 \le n < n_{eM}$, they will join. Otherwise, they will balk.

If the partial information is $\overline{s}_1$ and $\overline{s}_2$ (or the variance $\sigma_s^2 = \overline{s}_2 - \overline{s}_1^2$), then the equilibrium threshold is obviously obtained by solving $U_{join} = 0$ and it is independent of the actual service time distribution.

B. Risk-Averse Customers

Then we consider a system with risk-averse customers. In this case, any customer's residual utility of joining U_{join} can be expressed explicitly as

$$U_{join} = \int_0^\infty (R - ct^{m+1})dF(t), \quad \forall m \in N^+, \tag{2}$$

where m indicates the sensitive degree of the risk-averse customers to the potential joining risk. Obviously, customers are classified by their risk sensitivity that increases along with the value of m.

We assume the precondition that the actual distribution of the service time exists any order raw moment. If the arriving customers know the first l raw moments $\{\overline{s}_1, \overline{s}_2, \cdots, \overline{s}_l \mid 1 \le l \le m+1\}$ of the service time, where $E[S^k] = \overline{s}_k$, $1 \le k \le l$, then the following lemma shows the maximum entropy distribution of their sojourn time based on the partial information (see Kapur [22]).

Lemma 2 If risk-averse customers are informed of the first l ($1 \le l \le m+1$) raw moments $\{\overline{s}_1, \overline{s}_2, \cdots, \overline{s}_l\}$ of the service time in an observable queue, then the density function of maximum entropy distribution of their sojourn time is

$$f(t) = e^{-\lambda_0 - \lambda_1 t - \lambda_2 t^2 - \cdots - \lambda_l t^l}, \tag{3}$$

and $\lambda_0, \lambda_1, \lambda_2, \ldots, \lambda_l$ are the solutions of the following equations:

$$\begin{cases} e^{\lambda_0} = \int_0^\infty e^{-\lambda_1 t - \lambda_2 t^2 - \cdots - \lambda_l t^l} dt \\ E[T^k] e^{\lambda_0} \\ = \int_0^\infty t^k e^{-\lambda_1 t - \lambda_2 t^2 - \cdots - \lambda_l t^l} dt, \ k = 1, 2, \ldots, l \end{cases} \tag{4}$$

where (we denote $\overline{s}_0 = 1$)

$$\begin{aligned} & E[T^k] = E[(S_1^- + S_2 + \cdots + S_n)^k] \\ & = \sum_{\sum_{i=1}^n k_i = k} C_k^{k_1} C_{k-k_1}^{k_2} \cdots C_{k-\sum_{j=1}^{n-1} k_j}^{k_n} \times \\ & E\left[S_1^{-k_1} S_2^{k_2} \cdots S_n^{k_n}\right] \\ & = \sum_{\sum_{i=1}^n k_i = k} \frac{k!}{k_1!(k-k_1)!} \frac{(k-k_1)!}{k_2!(k-k_1-k_2)!} \\ & \cdots \times \frac{(k - \sum_{j=1}^{n-1} k_j)!}{k_n!(k - \sum_{j=1}^{n} k_j)!} \prod_{j=2}^n E\left[S_1^{-k_1}\right] E\left[S_j^{k_j}\right] \\ & = \sum_{\sum_{i=1}^n k_i = k} \frac{k!}{k_1! k_2! \cdots k_n!} \prod_{j=2}^n \overline{s}_{k_j} E[S_1^{-k_1}] \end{aligned} \tag{5}$$

and $E[S_1^{-k_1}]$ can be solved by differentiating the Laplace-Stieltjes transform (LST) of the residual service time $S^{-*}(s) = (1 - S^*(s))/(\overline{s}_1 s)$.

Theorem 2 If risk-averse customers are informed of the first l raw moments $\{\overline{s}_1, \overline{s}_2, \cdots, \overline{s}_l\}$ $(1 \le l \le m+1)$ of the service time in an observable queue, then the equilibrium threshold n_{elM} is the floor function of the solution of

$$R - c \int_0^\infty t^{m+1} e^{-\lambda_0 - \lambda_1 t - \lambda_2 t^2 - \cdots - \lambda_l t^l} dt = 0, \tag{6}$$

where $\lambda_0, \lambda_1, \lambda_2, \ldots, \lambda_l$ are the solutions of Eqs.(4) and (5).

III. Social Optimization

Now we respectively consider the socially optimal thresholds of risk-neutral and risk-averse customers if the actual service time distribution is exponential with parameter μ.

We first denote the expected social welfare per time unit by SW and socially optimal threshold by n^*. Given a maximum queue length n, the system is actually an M/M/1/n queue and the stationary state probabilities $\{p_i, i = 1, 2, \ldots, n\}$ are

$$p_i = \rho^i p_0, \ \rho = \frac{\lambda}{\mu}, \ i = 1, 2, \ldots, n,$$

where $p_0 = (1-\rho)/(1-\rho^{n+1})$. So the probability of observing n customers in the system, that is, the system loss probability, is

$$p_n = \rho^n p_0 = \frac{(1-\rho)\rho^n}{1-\rho^{n+1}}.$$

A. Risk-Neutral Customers

Given that the partial information is $\overline{s}_1 = 1/\mu$, based on Proposition 2.1, we have $n_{eM} = \lfloor R\mu/c \rfloor$. The social welfare per time unit, denoted by SW, is

$$\begin{aligned} & SW = \lambda(1-p_n)\left(R - c\int_0^\infty t dF(t)\right) \\ & = \lambda(1-p_n)\left(R - cE[T]\right) \\ & = \lambda(1-p_n)\left(R - \frac{c}{\mu(1-p_n)} \sum_{i=1}^n i p_{i-1}\right) \\ & = \lambda(1-p_n)\left(R - \frac{c}{\mu(1-p_n)} \sum_{i=0}^{n-1} (i+1) p_i\right) \\ & = \frac{\lambda(1-\rho^n)}{1-\rho^{n+1}} \left(R - c\left(\frac{1}{\mu(1-\rho)} - \frac{n\rho^n}{\mu(1-\rho^n)}\right)\right). \end{aligned} \tag{7}$$

So the socially optimal threshold n^* is the floor function of the maximizer of Eq.(7).

B. Risk-Averse Customers

Given risk-averse customers, the social welfare per time unit SW is

$$\begin{aligned} SW &= \lambda(1-p_n)\left(R - c\int_0^\infty t^{m+1} dF(t)\right) \\ &= \frac{\lambda(1-\rho^n)}{1-\rho^{n+1}} \left(R - cE[T^{m+1}]\right), \end{aligned} \tag{8}$$

where

$$\begin{aligned} & E[T^{m+1}] \\ & = \sum_{\sum_{i=1}^n k_i = m+1} \sum_{\sum_{i=1}^n k_i j_{k_i} = m+1} \prod_{k_i=1}^{m+1} \left(\frac{k_i!}{\mu^{k_i}}\right)^{j_{k_i}} \\ & = \frac{1}{\mu^{m+1}} \sum_{\sum_{i=1}^n k_i = m+1} \sum_{\sum_{i=1}^n k_i j_{k_i} = m+1} \prod_{k_i=1}^{m+1} (k_i!)^{j_{k_i}}. \end{aligned}$$

1) Special Case: $m = 1$:

- **Mean** Based on Lemma 2.1, given that the partial information is the mean $\overline{s}_1 = 1/\mu$ of the service time, then the maximum entropy distribution is exponential that just coincides with the actual distribution, that is, the customers' estimation is accurate. So the maximum entropy distribution of the sojourn time is Erlang distribution with parameter n and $1/\overline{s}_1$, then we get $U_{join} = R - cn(n+1)\overline{s}_1^2$. Solving $U_{join} = 0$, we get the equilibrium threshold as

$$n_{eM} = \left\lfloor -\frac{1}{2} + \sqrt{\frac{1}{4} + \frac{R\mu^2}{c}} \right\rfloor. \tag{9}$$

On the other hand, when $m = 1$ the expected social welfare per time unit is

$$\begin{aligned} & SW = \lambda(1-p_n)\left(R - c\int_0^\infty t^2 dF(t)\right) \\ & = \lambda(1-p_n)\left(R - cE[T^2]\right) \\ & = \lambda(1-p_n)\left(R - \frac{c}{\mu^2(1-p_n)} \sum_{i=0}^{n-1} (i+1)(i+2) p_i\right) \\ & = \frac{\lambda(1-\rho^n)}{1-\rho^{n+1}} \left(R - c\left(\frac{1}{\mu^2(1-\rho^n)} \left(\frac{2\rho^2(1-\rho^{n-2})}{(1-\rho)^2} \right.\right.\right. \\ & \left. -3n\rho^n + \frac{\rho + \rho^{n+1}(n-1)^2 - \rho^n(n^2-2)}{1-\rho}\right) \\ & \left.\left. + \frac{2+\rho}{\mu^2(1-\rho)}\right)\right), \end{aligned} \tag{10}$$

and the socially optimal threshold n^* is the floor function of the maximizer of Eq.(10).

- **Mean and Variance** Given that the partial information is both the mean $\bar{s}_1 = 1/\mu$ and variance $\sigma_s^2 = 1/\mu^2$ of the service time, then it is enough for the customers to make decisions so that the equilibrium threshold n_{eMV} is equal to that given in Eq.(9).

2) Special Case: $m = 2$:

- **Mean** Given that the partial information is $\bar{s}_1 = 1/\mu$ of the service time, we get $U_{join} = R - cn(n+1)(n+2)\bar{s}_1^3$. So the equilibrium threshold n_{eM} is the floor function of the solution of $U_{join} = 0$. On the other hand, when $m = 2$,

$$\begin{aligned} SW &= \lambda(1-p_n)\left(R - c\int_0^\infty t^3 dF(t)\right) \\ &= \lambda(1-p_n)\left(R - cE[T^3]\right) \\ &= \lambda(1-p_n)\left(R - \frac{c}{\mu^3(1-p_n)}\right. \\ &\quad \left.\times\sum_{i=0}^{n-1}(i+1)(i+2)(i+3)p_i\right) \\ &= \frac{\lambda(1-\rho^n)}{1-\rho^{n+1}}\left(R - c\left(\frac{1-\rho}{\mu^3(1-\rho^n)}\right.\right. \\ &\quad \times\left(S_{n-1}^{(3)} + 3S_{n-1}^{(2)}\right) \\ &\quad + \frac{2}{\mu^3(1-\rho^n)}\left(\frac{\rho(1-\rho^{n+1})}{1-\rho} - \rho^n(\rho+n)\right) \\ &\quad + \frac{3}{\mu^3(1-\rho^n)}\left(\frac{2\rho^2(1-\rho^{n-2})}{(1-\rho)^2}\right. \\ &\quad + \frac{\rho^{n+1}(n-1)^2 - \rho^n(n^2-2)}{1-\rho} - n\rho^n \\ &\quad \left. + \frac{\rho}{1-\rho}\right) + \frac{3\rho}{\mu^3(1-\rho)} \\ &\quad \left.\left. + \frac{6}{\mu^3}\left(\frac{1}{1-\rho} - \frac{n\rho^n}{1-\rho^n}\right)\right)\right), \end{aligned} \tag{11}$$

where

$$\begin{aligned} S_{n-1}^{(2)} &= \sum_{i=1}^{n-1} i^2\rho^i \\ &= \frac{\rho - \rho^n(n^2-2) + \rho^{n+1}(n-1)^2}{(1-\rho)^2} \\ &\quad + \frac{2\rho^2(1-\rho^{n-2})}{(1-\rho)^3} \end{aligned}$$

and

$$\begin{aligned} S_{n-1}^{(3)} &= \sum_{i=1}^{n-1} i^3\rho^i \\ &= \frac{2+\rho}{1-\rho}\left(\frac{\rho - \rho^n(n^2-2) + \rho^{n+1}(n-1)^2}{(1-\rho)^2}\right. \\ &\quad \left. + \frac{2\rho^2(1-\rho^{n-2})}{(1-\rho)^3}\right) - \left(\frac{\rho^2 - n\rho^n}{(1-\rho)^2}\right. \\ &\quad \left. + \frac{\rho^2(1-\rho^{n-1})}{(1-\rho)^3}\right) - \frac{n(n-1)^2\rho^n + \rho}{1-\rho}, \end{aligned}$$

the socially optimal threshold n^* is the floor function of the maximizer of Eq.(11).

- **Mean and Variance** When the actual service time distribution is exponential, maximum entropy distribution of the service time is also exponential given both the mean $\bar{s}_1 = 1/\mu$ and variance $\sigma_s^2 = 1/\mu^2$ (see Kapur [22]). So the equilibrium threshold n_{eMV} coincides with n_{eM} when $m = 2$.
- **First Three Raw Moments** Given that the partial information is the first three raw moments $\bar{s}_1, \bar{s}_2, \bar{s}_3$ of service time, then $\bar{s}_1 = 1/\mu, \bar{s}_2 = 2/\mu^2, \bar{s}_3 = 6/\mu^3$. Hence it is enough for the customers to make decisions so that the equilibrium threshold n_{e3M} equals to n_{eM} when $m = 2$.

IV. Profit Maximization

Besides the equilibrium and socially optimal thresholds, we also consider the profit-maximizing threshold denoted by n_m, i.e., the desired threshold chosen by the server. For the risk-neutral case, Hassin and Haviv [1] presented that the profit-maximizing threshold n_m is the floor function of the solution of equation

$$\nu + \frac{(1-\rho^{\nu-1})(1-\rho^{\nu+1})}{\rho^{\nu-1}(1-\rho)^2} - \frac{R\mu}{c} = 0.$$

Then we try to derive the profit-maximizing thresholds for two risk-averse cases and first consider $m = 1$ then $m = 2$. Conformed with the threshold n_m, the price set by the server, denoted by p_m, is

$$p_m = R - cE[T^2] = R - \frac{cn_m(n_m+1)}{\mu}.$$

So the server's profit, denoted by P_{n_m}, is

$$\begin{aligned} P_{n_m} &= \lambda(1-p_{n_m})p_m \\ &= \frac{\lambda(1-\rho_m^n)}{1-\rho^{n_m+1}}\left(R - \frac{cn_m(n_m+1)}{\mu}\right) \\ &= \frac{\lambda R(1-\rho_m^n)}{1-\rho^{n_m+1}}\frac{\nu_e(\nu_e+1) - n(n+1)}{\nu_e(\nu_e+1)}, \end{aligned} \tag{12}$$

where $\nu_e(\nu_e+1) = R\mu^2/c$. Because a profit-maximizing threshold satisfies the following two conditions: $P_{n_m} > P_{n_m-1}$ and $P_{n_m} \leq P_{n_m+1}$(see Hassin [1]). Substituting Eq.(12) into the first condition, then it amounts to

$$\nu_e(\nu_e+1) > f_1(n_m),$$

where

$$f_1(n_m) = n_m(n_m+1) + \frac{2n_m(1-\rho^{n_m-1})(1-\rho^{n_m+1})}{\rho^{n_m-1}(1-\rho)^2}.$$

Substituting $n_m + 1$ for n_m and reversing direction of the inequality, the second condition becomes $\nu_e(\nu_e+1) \leq f_1(n_m+1)$, where

$$\begin{aligned} f_1(n_m+1) &= (n_m+1)(n_m+2) \\ &\quad + \frac{2(n_m+1)(1-\rho^{n_m})(1-\rho^{n_m+2})}{\rho^{n_m}(1-\rho)^2}. \end{aligned}$$

Then the two conditions can be summarized to

$$f_1(n_m) < \nu_e(\nu_e+1) \leq f_1(n_m+1).$$

Define a function

$$\begin{aligned} f_1(\nu) &= \nu(\nu+1) \\ &\quad + \frac{2\nu(1-\rho^{\nu-1})(1-\rho^{\nu+1})}{\rho^{\nu-1}(1-\rho)^2} - \nu_e(\nu_e+1), \end{aligned} \tag{13}$$

which monotonously increases from 0 to ∞ with ν. So there exists a unique solution ν_m^1 to $f_1(\nu) = 0$ and then $n_m = \lfloor\nu_m^1\rfloor$ when $m = 1$.

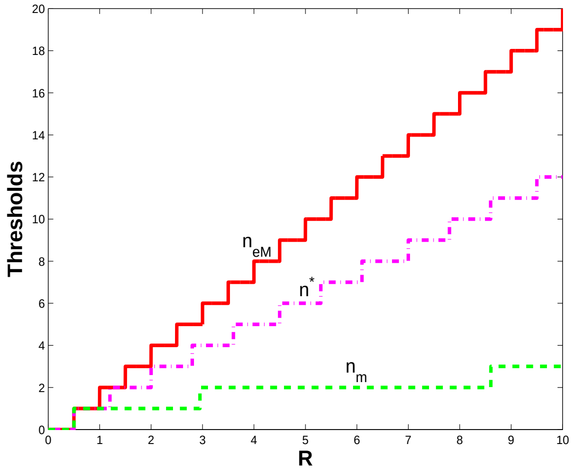


Fig. 1. Comparison of thresholds when $m = 0$

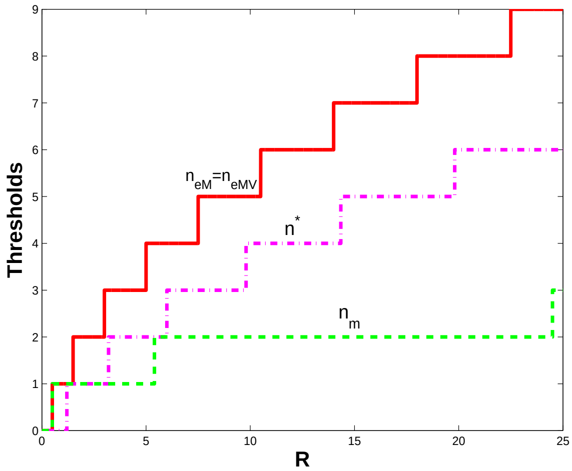


Fig. 2. Comparison of thresholds when $m = 1$

Similar to the risk-averse case $m = 1$, when $m = 2$, the price set by the server is

$$p_m = R - cE[T^3] = R - \frac{cn_m(n_m+1)(n_m+2)}{\mu}.$$

So the server's profit is

$$\begin{aligned} P_{n_m} &= \frac{\lambda R(1-\rho_m^n)}{1-\rho^{n_m+1}} \\ &\times \frac{\nu_e(\nu_e+1)(\nu_e+2) - n(n+1)(n+2)}{\nu_e(\nu_e+1)(\nu_e+2)}, \end{aligned} \tag{14}$$

where $\nu_e(\nu_e+1)(\nu_e+2) = R\mu^3/c$. After simplicities and based on the same two conditions, we define another function

$$\begin{aligned} & f_2(\nu) = \nu(\nu+1)(\nu+2) \\ & + \frac{3\nu(\nu+1)(1-\rho^{\nu-1})(1-\rho^{\nu+1})}{\rho^{\nu-1}(1-\rho)^2} - \frac{R\mu^3}{c}, \end{aligned} \tag{15}$$

which also monotonously increases from 0 to ∞ with ν. Hence there exists a unique solution ν_m^2 to $f_2(\nu) = 0$ and then $n_m = \lfloor \nu_m^2 \rfloor$ when $m = 2$.

Figs. 1–3 numerically make comparisons of the equilibrium threshold(s) n_e and socially optimal threshold n^* as well as profit-maximizing threshold n_m when $m = 0, 1, 2$ respectively. They jointly show that the three types of thresholds all decrease along with the increase of m.

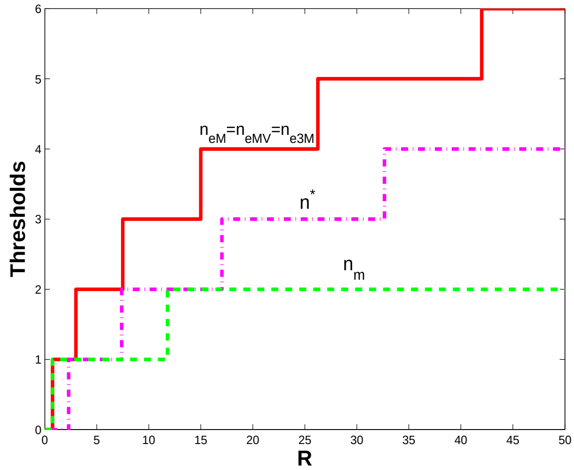


Fig. 3. Comparison of thresholds when $m = 2$

Moreover, Fig. 1 verifies the relationship pointed out by Naor [23] that $n_{eM} \geq n^* \geq n_m$ for the risk-neutral case. However, Figs. 2 and 3 show that for the risk-averse case the relationship $n_{eM} = n_m > n^*$ is still possible for the relatively smaller values of R. In addition, because of the particularity of exponential distribution, we illustrate in Fig. 3 that $n_{e3M} = n_{eMV} = n_{eM}$, i.e., the equilibrium thresholds under the three types of partial information are all equal given that the actual service time distribution is exponential.

V. Conclusion and Future Work

Based on the maximum entropy principle and partial service information, we considered customer threshold strategies of joining or balking in an observable queue. Regardless of risk-neutral and risk-averse customers, we observed that the equilibrium thresholds are no less than the socially optimal and profit-maximizing ones, and more risk-averse customers hold lower threshold. What is interested for us is the following question: Which is the minimax regret distribution of the service time when partial information is the mean or mean and variance? Considering the minimax regret criterion is the most humanistic decision criterion, this will be our next step work.

References

[1] R. Hassin, and M. Haviv, Equilibrium Behavior in Queueing Systems: To Queue or not to Queue, Kluwer Academic Publishers, Boston, 2003.

[2] P. Guo, and P. Zipkin, "Analysis and comparison of queues with different levels of delay information," *Management Science*, vol. 53, no. 6, pp. 962-970, 2007.

[3] P. Guo, and P. Zipkin, "The effects of information on a queue with balking and phase-type service times," *Naval Research of Logistics*, vol. 55, pp. 406-411, 2008.

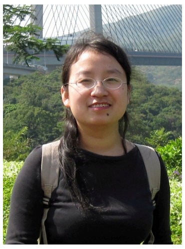
[4] P. Guo, and P. Zipkin, "The effect of the availability of waiting-time information on a balking queue," *European Journal of Operational Research*, vol. 198, pp. 199-209, 2009.

[5] R. Ibrahim, and W. Whitt, "Real-time delay estimation based on delay history," *Manufacturing Service Operations Management*, vol. 11, no. 3, pp. 397-415, 2009.

[6] R. Ibrahim, and W. Whitt, "Real-Time delay estimation in overloaded multiserver queues with abandonments," *Management Science*, vol. 55, no. 10, pp. 1729-1742, 2009.

[7] M. Armony, N. Shimkin, and W. Whitt, "The impact of delay announcements in many-server queues with abandonment," *Operations Research*, vol. 57, no. 1, pp. 66-81, 2009.

[8] O. Jouini, Y. Dallery, and O.Z. Aksin, "Queueing models for multiclass call centers with real-time anticipated delays," *International Journal of Production Economics*, vol. 120, pp. 389-399, 2009.
[9] Y. Chen, P. Chen, and Y. Zhu, "Analysis of a call center with partial closing rules, feedback and impatient calls," *IAENG International Journal of Applied Mathematics*, vol. 46, no. 4, pp. 585-591, 2016.
[10] W. Sun, P. Guo, and N. Tian, "Equilibrium threshold strategies in observable queueing systems with setup/closedown times," *Central European Journal of Operations Research*, vol. 18, no. 3, pp. 241-268, 2010.
[11] W. Sun, Y. Wang, and N. Tian, "Pricing and setup/closedown policies in unobservable queues with strategic customers," *4OR: A Quarterly Journal of Operations Research*, vol. 10, no. 3, pp. 287-311, 2012.
[12] P. Guo, and R. Hassin, "Strategic behavior and social optimization in Markovian vacation queues. Operations Research, vol. 59, no. 4, pp. 986-997, 2011.
[13] P. Guo, and R. Hassin, "Strategic behavior and social optimization in Markovian vacation queues: The case of heterogeneous customers. European Journal of Operational Research, vol. 222, no. 2, pp. 278-286, 2012.
[14] A. Economou, A. Gómez-Corral, and S. Kanta, "Optimal balking strategies in single-server queues with general service and vacation times," *Performance Evaluation*, vol. 68, no. 10, pp. 967-982, 2011.
[15] W. Sun, S. Li, and N. Tian, "Equilibrium mixed strategies of customers in an unobservable queue with multiple vacations," *Quality Technology & Quantitative Management*, vol. 10, no. 4, pp. 389-421, 2013.
[16] W. Sun, and S. Li, "Equilibrium and optimal behavior of customers inMarkovian queues with multiple working vacations," *TOP*, vol. 22, no. 2, pp. 694-715, 2014.
[17] W. Sun, S. Li, and Q. Li, "Equilibrium balking strategies of customers in Markovian queues with two-stage working vacations," *Applied Mathematics and Computation*, vol. 248, pp. 195-214, 2014.
[18] F. Wang, J. Wang, and F. Zhang, "Equilibrium customer strategies in the Geo/Geo/1 queue with single working vacation," *Discrete Dynamics in Nature and Society*, vol. 2014, pp. 185-191, 2014.
[19] G. Panda, V. Goswami, and A.D. Banik, "Equilibrium and socially optimal balking strategies in Marko- vian queues with vacations and sequential abandonment," *Asia-Pacific Journal of Operational Research*, vol. 33, no. 5, pp. 1-34, 2016.
[20] W. Sun, S. Li, and N. Tian, "Equilibrium and optimal balking strategies of customers in unobservable queues with double adaptive working vacations," *Quality Technology and Quantitative Management*, vol. 14, no. 1, pp. 94-113, 2017.
[21] P. Guo, W. Sun, and Y. Wang, "Equilibrium and Optimal Strategies to Join a Queue with Partial Information on Service Times," *European Journal of Operational Research*, vol. 214, no. 2, pp. 284-297, 2011.
[22] J.N. Kapur, and H.K. Kesavan, Entropy optimization principles with applications, Academic Press, Boston, 1992.
[23] P. Naor, "The regulation of queue size by levying tolls," *Econometrica*, vol. 37, pp. 15-24, 1969.



Wei Sun received her Ph.D. degree from Yanshan University, Qinhuangdao, China, in 2010. Currently she is an associate professor in the School of Economics and Management at Yanshan University. She has published more than 40 papers in international leading journals in the areas of operations research, and applied mathematics. She has been involved in several projects supported by the National Natural Science Foundation of China, the National Education Committee Foundation of China, and other foundations. Her research interests include economics of queues, and queueing systems with vacations.